

Differential Privacy: Messaging and Communication Strategies for General Audience

Discussant Comments

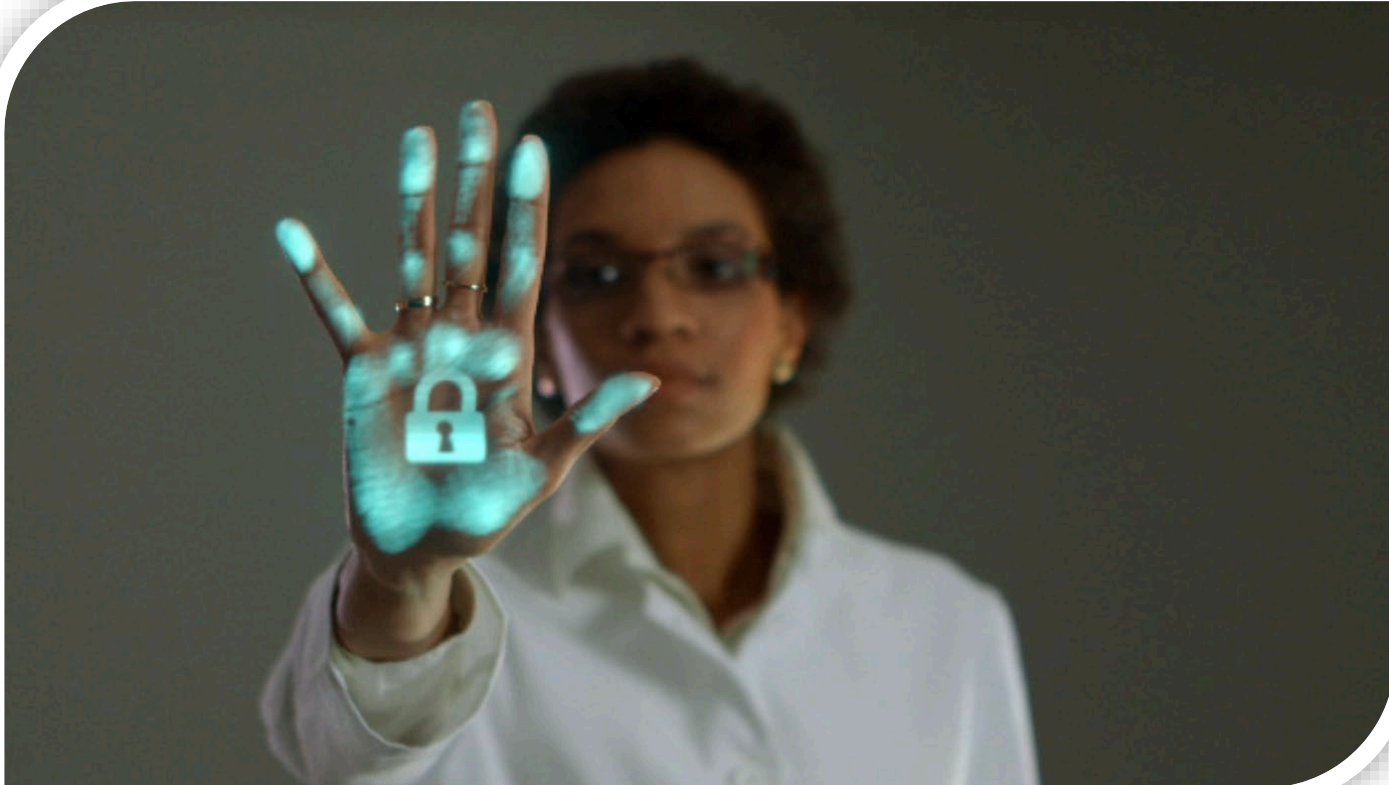
By Julio Guity-Guevara | John Sandoval

Census National Advisory Committee

Implementing Differential Privacy for the 2020 Census Data Products Working Group

March 7, 2022

Background: Impacts for Racial and ethnic populations of Previous Interpretations of Census Privacy Laws

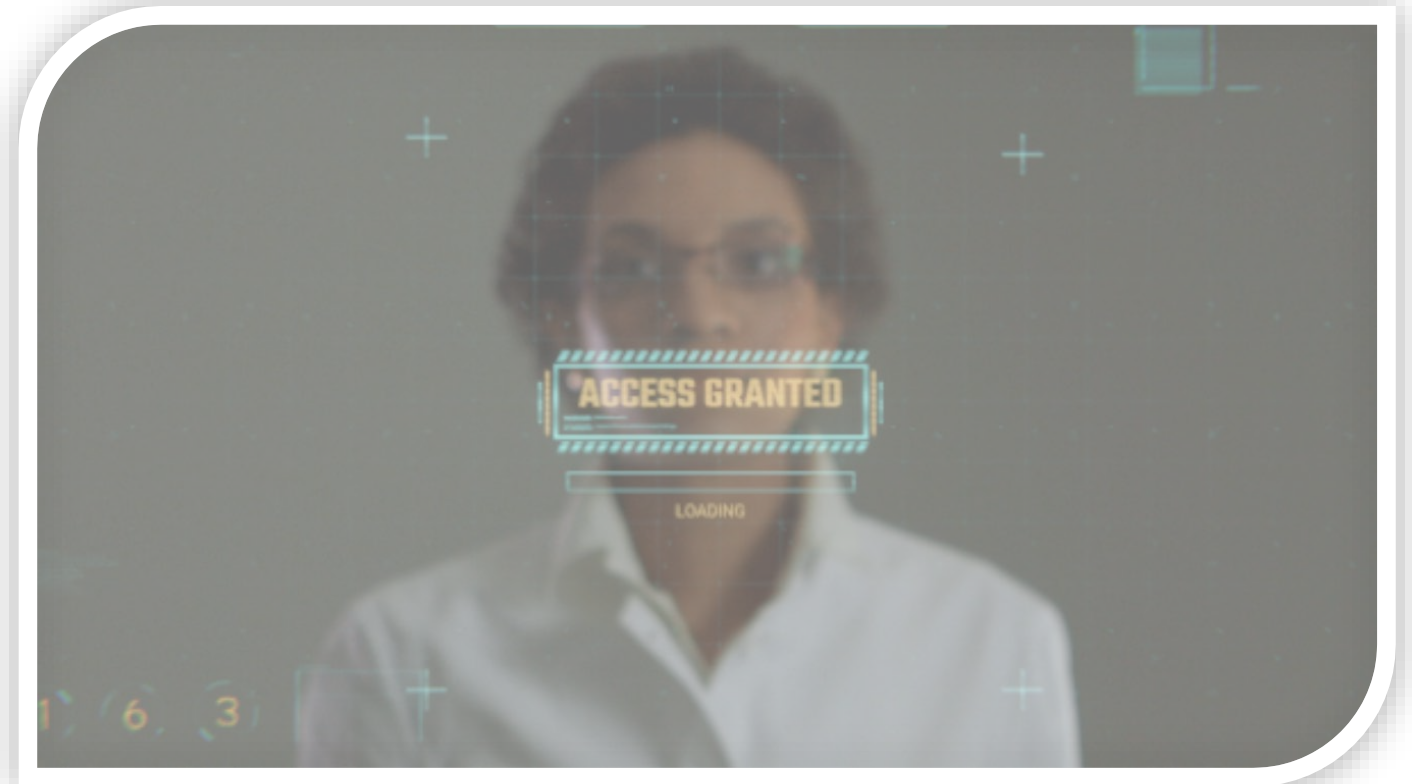
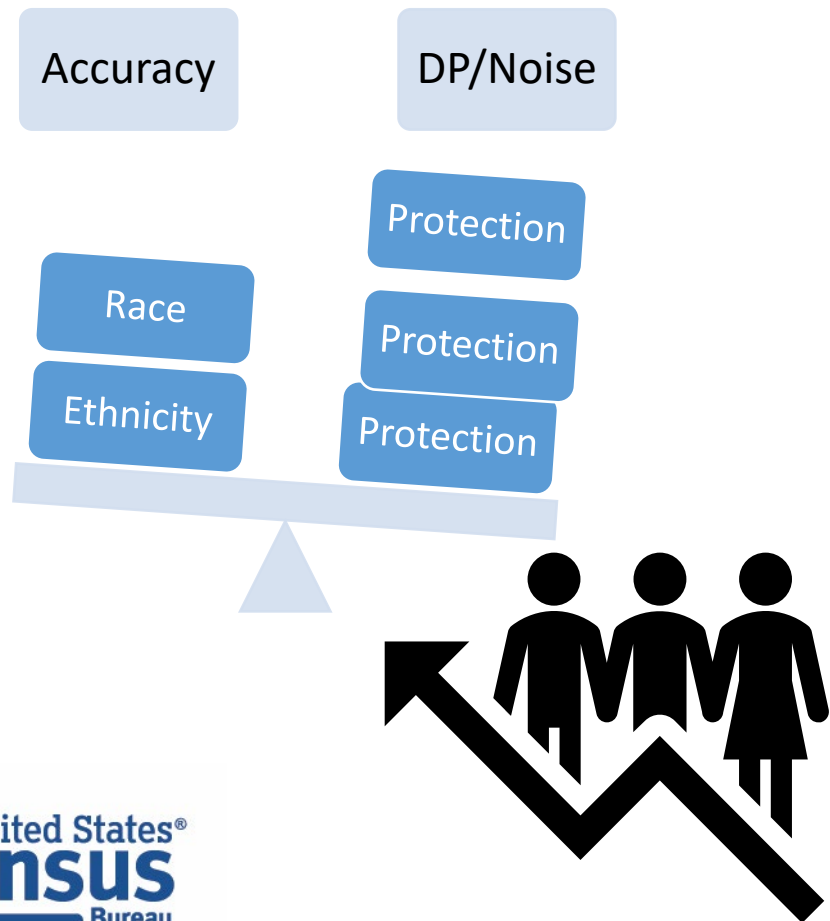


Title 13 of the United States Code



Impacts for Historically Undercounted Populations

Race and Ethnic Data Quality



Census Bureau has proactively shared information about DAS and Differential Privacy

United States
Census
Bureau

Search

BROWSE BY TOPIC

EXPLORE DATA

LIBRARY

SURVEYS/ PROGRAMS

INFORMATION FOR...

FIND A CODE

2020 Census Data Quality Results to be released March 10

// [Census.gov](#) / [2020 Census Program Management](#) / [Processing the Count](#) / [Disclosure Avoidance Modernization](#) / [Understanding Differential Privacy](#)

WITHIN DECADENNAL CENSUS OF POPULATION AND HOUSING

About

By Decade

Data

Geographies

Guidance for Data Users

Information for Respondents

Library

Newsroom

Technical Documentation

Contact Us

f

t

in

Understanding Differential Privacy

Census confidentiality protections—what we call “disclosure avoidance”—have evolved over time to keep pace with emerging threats. Since the 1990 Census we’ve added “noise”—or variations from the actual count—to the collected data. For 2020 Census data we’re applying noise using a newer protection framework based on “differential privacy.” Learn more here about why and how we’re modernizing our protections and how you can engage in the process.

On this page:

- [Why We’re Modernizing Census Disclosure Avoidance](#)
- [How Census Disclosure Avoidance Methods Have Evolved Over Time](#)
- [How Differential Privacy Works: The Basics](#)
- [Technical Background Information](#)

Why We’re Modernizing Census Disclosure Avoidance

Frequently Asked Questions

[What law requires the Census Bureau to protect against disclosure of my Census responses?](#)

[Why do we need a different disclosure avoidance system in 2020 than we used in 2010?](#)

[What harms can arise if the basic demographic data collected in the decennial census are exposed?](#)

[Could external attackers know whether they’ve correctly reidentified individuals in Census data even if they don’t have access to confidential census records?](#)

Comparing Differential Privacy With Order Disclosure Avoidance Methods

The U.S. Census Bureau's methods to protect your responses in published census data have evolved steadily over the decades. In the past century, we moved from a system that relied on "eyeballing" data tables to spot potentially revealing statistics to a system of intricate, statistical techniques to address group disclosure risk. But the methods used only people in their block with a specific combination of sex, age (in years), race (any of the 63 possible Office of Management and Budget race combinations), and Hispanic/Latino ethnicity.¹ Those kinds of unique attributes are precisely the vulnerabilities discoverable by today's technology.

The decision to adopt confidentiality protections based on differential privacy for the 2020 Census was based on research that exposed the limits of our previous methods. We conducted experiments to better understand how those techniques would impact census results if applied today, would publish 2020 Census results at the block.

The findings offer stakeholders a tool for comparing the trade-offs between those earlier methods and the new approach designed for publication in the P.L. 94-171 redistricting data, the TopDown Algorithm, which is based on the principles of differential privacy.


These findings are summarized below and are also available via a webinar we conducted in June 2021 at www.census.gov/data/jacemy/webinars/2021/disclosure-avoidance-series/research-on-alternatives-to-differential-privacy.html.

DIFFERENTIAL PRIVACY

What Is Differential Privacy and How Does It Work?

Differential privacy, first developed in 2006, is a framework for measuring the precise disclosure risk associated with each release of confidential data. It allows an agency like the Census Bureau to quantify the precise amount of statistical noise required to protect confidentiality. This precision allows us to calibrate and allocate precise amounts of statistical noise in a way that protects confidentiality while maintaining the overall accuracy of the data in the aggregate.

The amount of randomly generated noise that is injected is driven by a tunable, or adjustable, "privacy-loss budget." An algorithm, that is also tunable, determines how much of that noise is injected into individual results and geographies. It is important to note that, since publishing exact census data for housing units at low levels of geography is the key to re-identifying the people behind the statistics, the new design disclosure avoidance system limits the kinds of statistics that are published as counted. These are called invariants. The 2020 Census publishes exact counts for the total population at the state level, the number and type of occupied group quarters facilities at the block level, and the number of housing units, whether occupied or not, at the block level.



1. Approved for public release per Declass Review Board response number: C2006P-7723-025-003

DP-14-0018 (Rev. August 2021)

WITHIN CENSUS BUREAU

Courses

Data Gems

Request a Training

Resources

Webinars

f

Twitter

in

Differential Privacy 101

MAY 04, 2021

Description

This introductory webinar will explain the Census Bureau's decision to modernize the methods we use to protect respondent confidentiality in the data we release. We will cover the basic principles and concepts of disclosure avoidance and differential privacy and discuss their implications for census data users.

Accuracy/Privacy Tradeoff

100%

Accuracy

0

private, inaccurate

1 out of the hat. In general, the idea is this: more privacy means you get less accuracy. Less

Protecting Privacy with MATH (Collab with the Census)

406,231 views • Sep 12, 2019

1BK

DISLIKE

SHARE

DOWNLOAD

THANKS

CLIP

SAVE

...

WITHIN CENSUS BUREAU

Comment Policy

Director's Blog

Global Reach

Random Samplings

Research Matters

f

Twitter

in

Modernizing Privacy Protections for the 2020 Census: Next Steps

April 28, 2021

DR. JOHNNIE M. ABOU-D, CHIEF SCIENTIST AND ASSOCIATE DIRECTOR FOR RESEARCH AND METHODOLOGY AND DR. VICTORIA A. VELKOFF, ASSOCIATE DIRECTOR FOR DEMOGRAPHIC PROGRAMS

A HISTORY OF CENSUS PRIVACY PROTECTIONS

From 1790 to 2020, the U.S. Census Bureau has been a leader in protecting the privacy of the people it serves. This infographic highlights key milestones in the history of census privacy protections.

1790-1840: The first U.S. Census was conducted in 1790. The Census Act of 1790 established the Census Bureau and required that census data be kept confidential.

1850-1890: The Census Act of 1850 was the first to explicitly state that census data was to be kept confidential. The Census Act of 1880 further strengthened these protections.

1900-1940: The Census Act of 1900 was the first to require that census data be kept confidential for 72 years after the census was conducted. The Census Act of 1929 extended this protection to 100 years.

1950-1990: The Census Act of 1950 was the first to require that census data be kept confidential for 100 years after the census was conducted. The Census Act of 1976 extended this protection to 120 years.

2000-2020: The Census Act of 2002 was the first to require that census data be kept confidential for 120 years after the census was conducted. The Census Act of 2012 extended this protection to 150 years.

Key Milestones:

- 1790:** First U.S. Census
- 1850:** Census Act of 1850
- 1880:** Census Act of 1880
- 1900:** Census Act of 1900
- 1929:** Census Act of 1929
- 1950:** Census Act of 1950
- 1976:** Census Act of 1976
- 2002:** Census Act of 2002
- 2012:** Census Act of 2012

Challenges:

- 1950s:** The U.S. Supreme Court ruled in *Skidmore v. Flint* that the Census Bureau's privacy protections were not sufficient.
- 1970s:** The U.S. Supreme Court ruled in *United States v. Egan* that the Census Bureau's privacy protections were not sufficient.
- 1980s:** The U.S. Supreme Court ruled in *United States v. Egan* that the Census Bureau's privacy protections were not sufficient.
- 1990s:** The U.S. Supreme Court ruled in *United States v. Egan* that the Census Bureau's privacy protections were not sufficient.
- 2000s:** The U.S. Supreme Court ruled in *United States v. Egan* that the Census Bureau's privacy protections were not sufficient.
- 2010s:** The U.S. Supreme Court ruled in *United States v. Egan* that the Census Bureau's privacy protections were not sufficient.
- 2020s:** The U.S. Supreme Court ruled in *United States v. Egan* that the Census Bureau's privacy protections were not sufficient.

Conclusion: The U.S. Census Bureau has a long history of protecting the privacy of the people it serves. This infographic highlights key milestones in the history of census privacy protections.

Reporting on Differential Privacy highlights tradeoff between accuracy and privacy

The Washington Post
Democracy Dies in Darkness

Social Issues

New system to protect census data may compromise accuracy, some experts say

By Tara Bahrapour and Marissa J. Lang

June 1, 2021 at 7:05 p.m. EDT

ARIZONA

The census has a new process to protect your privacy. It also risks a less accurate count



Geoff Hing

Arizona Republic

Published 8:30 a.m. MT Aug. 10, 2021 | Updated 3:45 p.m. MT Aug. 10, 2021

Los Angeles Times

WORLD & NATION

Judges hear arguments over census privacy tool

Bloomberg CityLab

Sign In

Data Scientists Square Off Over Trust and Privacy in 2020 Census

As the Census Bureau releases redistricting data, researchers are split over a security measure called differential privacy that can blur population figures.

The New York Times

• **TheUpshot**

To Reduce Privacy Risks, the Census Plans to Report Less Accurate Data

Guaranteeing people's confidentiality has become more of a challenge, but some scholars worry that the new system will impede research.

Some media outlets explicitly showcase the potential negative impacts of Differential Privacy



US & WORLD

The 2020 U.S. census is about to drop. But a new government privacy measure could mean chunks of the data are useless



Nami Sumida

Aug. 10, 2021 | Updated: Aug. 11, 2021 6:46 a.m.



URBAN
WIRE

DATA/
VIZ

FEATURES

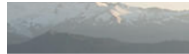
POLICY
DEBATES

PARTNER
PROJECTS

NEXT50

Urban Wire :: Research Methods and Data Analytics

The blog of the Urban Institute



March 4, 2020

Will the Census's Data Privacy Efforts Erase Rural America?



t TRUTHOUT

LATEST ABOUT DONATE

NEWS | RACIAL JUSTICE

New Census Algorithm Grossly Miscalculates People of Color in the Name of "Privacy"



AP

U.S. News World News Politics Sports Entertainment Business Technology Health Science Oddities Lifestyle

People, homes vanish due to 2020 census' new privacy method

By MIKE SCHNEIDER October 31, 2021

Communication Best Practices for General Audience

- **Opt for Nontechnical language – plain “English”**
- **Develop a 250-word definition of Differential Privacy that can be digested in one minute**
- **Use images, pictures and icons to facilitate understanding**
- **Stand alone document preferred over several inter-dependent documents**
- **Translate into the 59 languages and dialects used in 2020 Census**
- **Craft culturally appropriate messaging that has been vetted by in culture experts**
- **Work with key partners and stakeholders to deliver message and be open to feedback**

Crystal Clarity of Messaging around Title 13 and Differential Privacy is needed

Clarification

What does Title 13 require?

Why is Differential Privacy the best tool for the job?

What other tools were considered?

Why was DP not used before?

Is Differential Privacy moot when similar data is available from 3rd party sources?

John Abowd, chief scientist at the census, thinks the risk could grow, and argues that in any case, the bureau has a legal obligation to put the algorithm to use in 2020: “We’re already under a statutory mandate to enforce confidentiality protection,” he said. “We don’t have the luxury of rolling it out after we get it perfected.”

Will Differential Privacy Work for the 2020 Census?

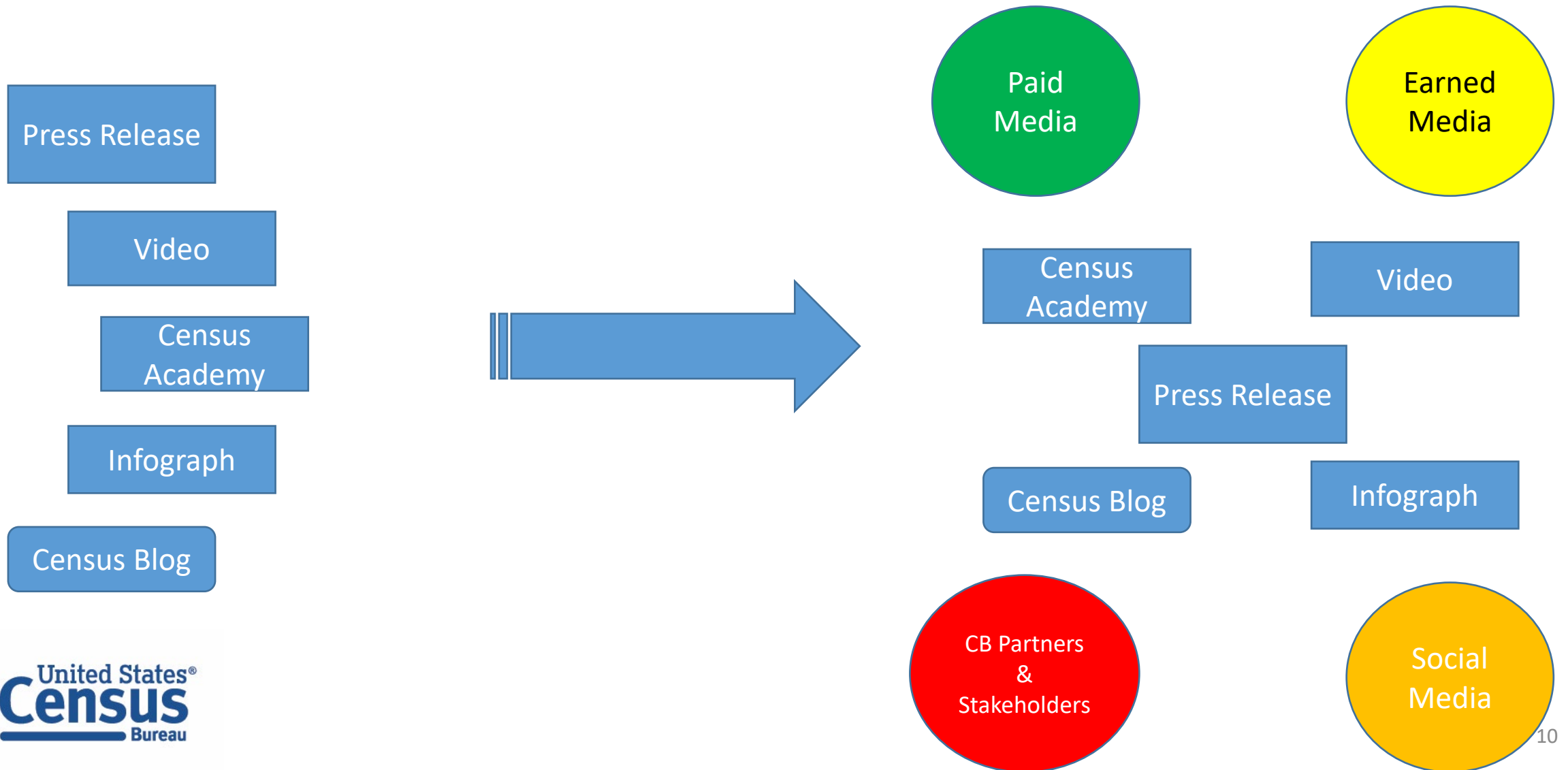
Yes. Differential privacy is the only framework that can provably protect 2020 Census data against known and emerging re-identification threats while producing quality, fit-for-use data. More information is available at <https://www.census.gov/library/video/2021/protecting-privacy-in-census-bureau-statistics.html>.

Defining WHAT the Message is requires more effort



- **Work with CB Partners and Stakeholders to uncover gaps in understanding and misconceptions**
- **Develop FAQs to bust myths, address doubts and help public grasp a complex topic**
- **Address concerns on Census Data Products Accuracy head-on**
- **Explain how Differential Privacy is a “Work In Process” and will be refined over time and use**
- **Share candid assessment of the impact of Differential Privacy on hard to count populations**
- **Reinforce value of participating in the Census and the accuracy and reliability of data and data products**

HOW you deliver the message is crucial



Differential Privacy

Considerations for Messaging the General Public

Bureau's decision to use differential privacy as part of its disclosure avoidance

Technological and Methodological Options

Outreach Strategies and Surveys



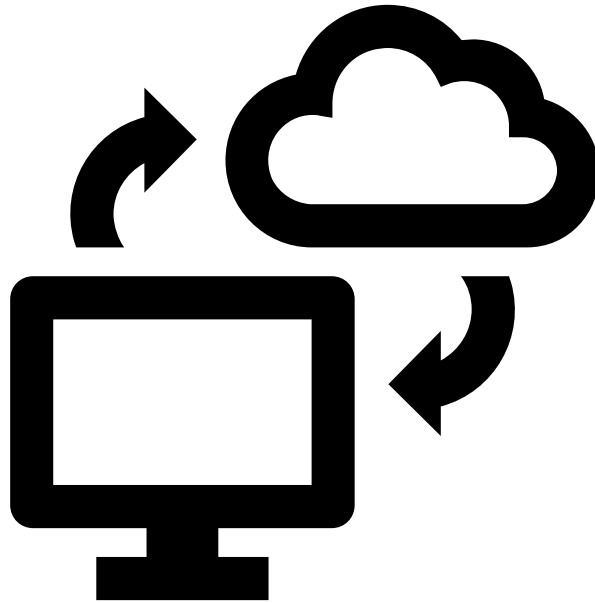
Differential Privacy

Considerations for Messaging the General Public

Bureau's decision to use differential privacy as part of its disclosure avoidance

Technological and Methodological Options

Outreach Strategies and Surveys



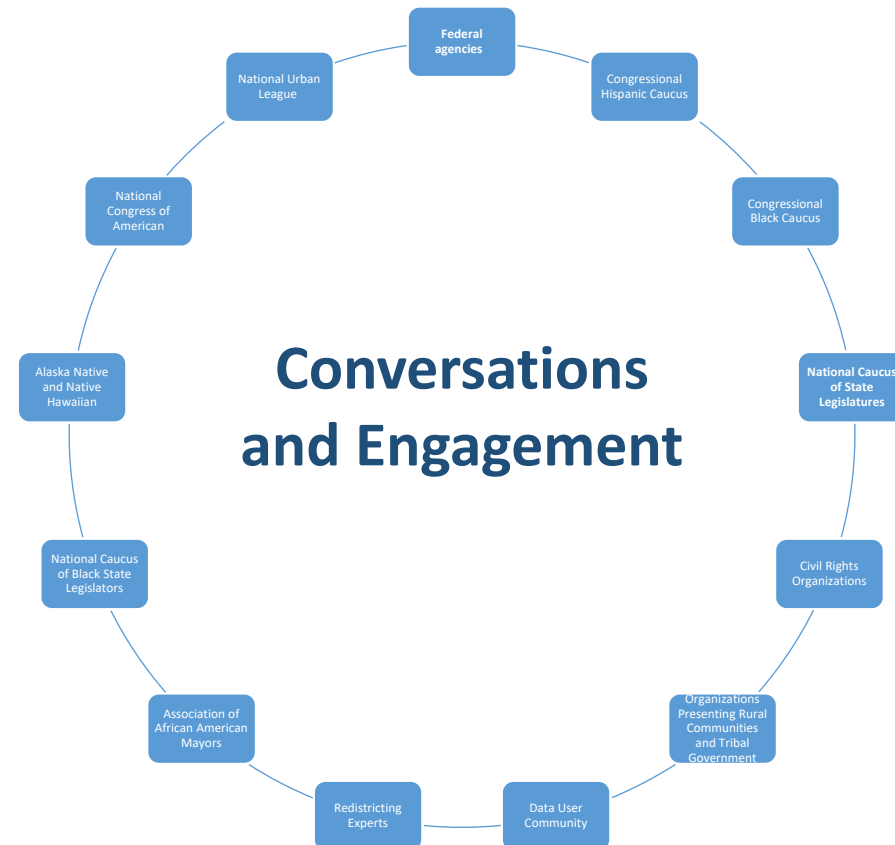
Differential Privacy

Considerations for Messaging the General Public

Bureau's decision to use differential privacy as part of its disclosure avoidance

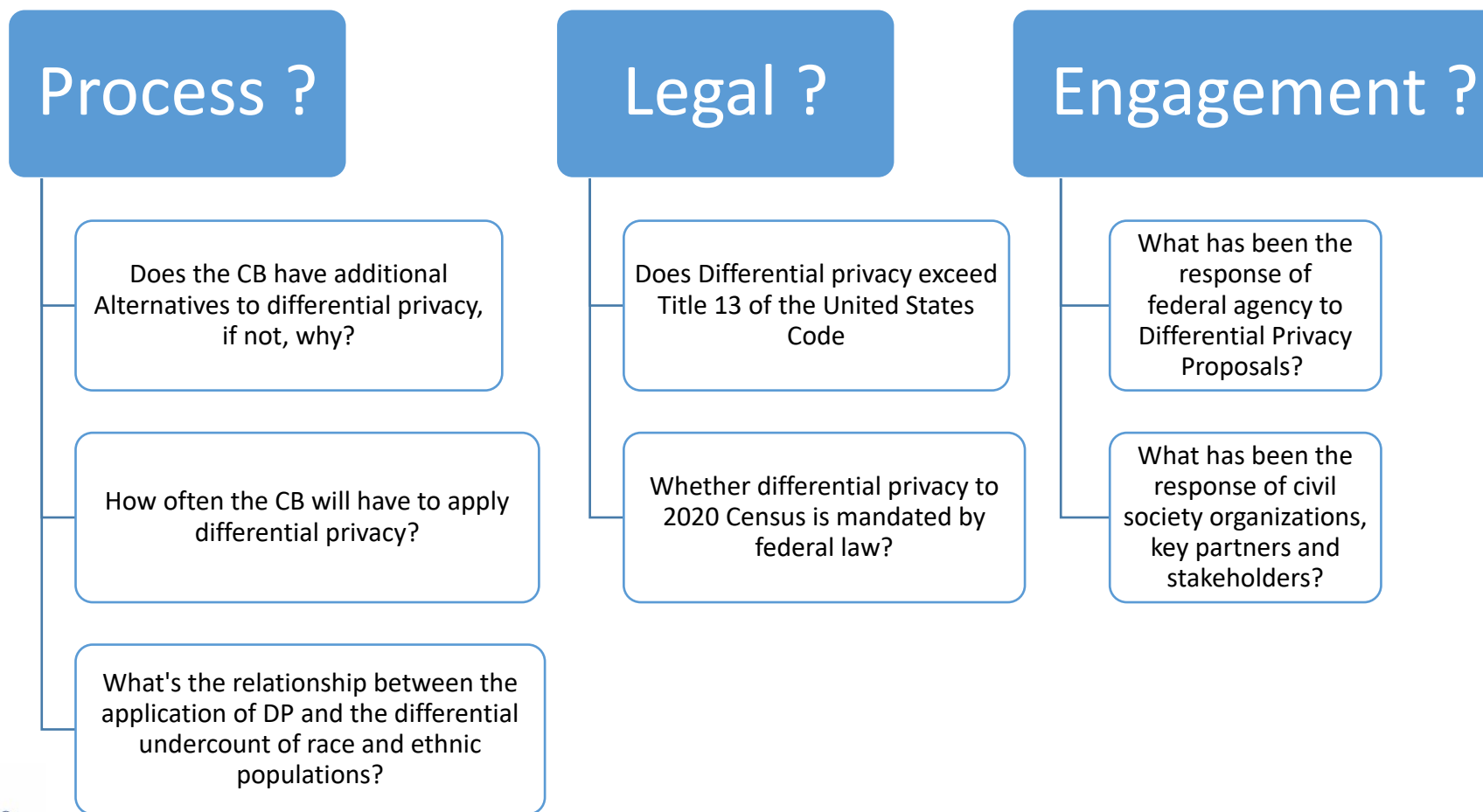
Technological and Methodological Options

Outreach Strategies and Surveys



Differential Privacy

Considerations for Messaging the General Public



COMMITTEE DISCUSSION

For audio to this conference, please dial 1-888-566-6192 (Access Code: 4881978#)